AD_____

GRANT NUMBER DAMD17-98-1-8335

TITLE: A Potential Tumor Suppressor Protein: Expression and Function in Human Breast Cancer Cells

PRINCIPAL INVESTIGATOR: Michelle J. Callaghan, Ph.D.

CONTRACTING ORGANIZATION: Garvan Institute of Medical Research
Darlinghurst NSW 2010 Australia

REPORT DATE: July 1999

TYPE OF REPORT: Annual Summary

PREPARED FOR: Commanding General
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1999	(1 Jul 98 - 30 Jun 99) Annual Summary

4. TITLE AND SUBTITLE
A Potential Tumor Suppressor Protein: Expression and Function in Human Breast Cancer Cells

5. FUNDING NUMBERS
DAMD17-98-1-8335

6. AUTHOR(S)
Michelle J. Callaghan, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Garvan Institute of Medical Research
Darlinghurst NSW 2010 Australia

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In a search for genes with a role in hormonal control of cell proliferation, we cloned a novel progestin-induced gene, EDD (previously DD5) that appears to be the human homologue of the *hyd* tumor suppressor gene of *Drosophila melanogaster*. By amino acid homology with other known proteins it is likely that the EDD protein is an E3 ubiquitin-protein ligase, enzymes which target one or more key proteins for destruction by ubiquitin-mediated proteolysis. During the first year of this project EDD was shown to be capable of E3 activity and antibodies useful in detecting and characterising EDD were obtained. To further define the cellular role of EDD eighteen candidate proteins were tested for interaction with EDD by immunoprecipitation. Nine proteins were tested using the yeast two-hybrid system. Yeast two-hybrid screening of cDNA libraries identified eight possible ubiquitination targets for EDD. So far two of these interactions have been confirmed by independent means. Studies of EDD in HEK-293 cells led to the discovery of an additional cellular role for EDD, that of a co-activator for progesterone mediated transcription. Finally experiments to define a possible role for EDD in human cancer have provided exciting data suggesting involvement of the EDD gene in several tumor types.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer Tumor suppressor genes, Steroid hormones/receptors	23 16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

_✓_ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46. N/A.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

M J Hardman 28th July 1999

PI - Signature Date

TABLE OF CONTENTS

ANNUAL SUMMARY- Year 1

AWARD NUMBER: DAMD17-98-1-8335

INTRODUCTION

In a search for genes with a role in hormonal control of cell proliferation, we have cloned a novel progestin-induced gene, EDD (previous designation DD5). It is hypothesized that EDD has a role in the control of cell growth and differentiation as it appears to be the human homologue of the *hyperplastic discs* (*hyd*) gene of *Drosophila melanogaster*. When mutated, *hyd* is either lethal or causes abnormal or hyperplastic growth in *Drosophila* larval imaginal discs and defective germ cell development. While the biochemical roles of *hyd* and EDD are unknown, by amino acid homology with other known proteins it is likely that the EDD protein is a ubiquitin-protein ligase (E3), enzymes which target one or more key proteins for destruction by ubiquitin-mediated proteolysis. The target proteins of EDD would be expected to have profound effects on cell cycle control or cell signalling. The project is designed to :

- define the normal function of EDD and its targets and relate this role to development and progression of breast cancer;
- determine the effects of EDD on breast cancer cell cycle progression;
- determine the role of EDD in progestin-induced growth stimulation of breast cancer cells; and
- determine the effect of mutations in EDD or dysregulation of EDD expression on tumor phenotype and tumor progression.

Over the first 12 months of this project significant progress has been made against these work objectives, including publication of evidence that EDD is indeed an E3 ligase, identification and testing of potential substrates for EDD, the generation of stable cell lines overexpressing EDD or a ubiquitination defective mutant form of EDD and acquisition of data supporting involvement of the EDD locus in a range of human tumor types. In addition to the proposed objectives, a new and exciting potential role has emerged for EDD.

BODY OF REPORT

TASK 1: To determine the ability of EDD protein to form a thioester bond with ubiquitin *in vitro* (months 1-6)

A key component of this objective involved generation of antisera to use in the detection and immunoprecipitation of EDD. Three of four peptide antibodies were found to be useful in this regard. Previously only an antibody raised to the *Drosophila* HYD protein was available and this antibody recognised a 100 kDa protein in human cells. Use of the EDD specific antibodies showed that the actual size of EDD was in fact around 300 kDa, as expected from the open reading frame within the cDNA and the reported size of the *Drosophila* protein (Fig 1; Callaghan et al, 1998). The EDD antibodies can be used to immunoprecipitate *in vitro* translated EDD protein, endogenous EDD protein from T-47D cell lysates or recombinant EDD protein from transiently transfected HEK-293 cells. One of the antibodies can be used to detect EDD protein by Western blotting.

Full-length EDD was *in vitro* transcribed and translated in the presence of GST-ubiquitin and shown to bind ubiquitin in a reversible manner. This bond was sensitive to reducing conditions, a characteristic of a thioester bond (Callaghan et al, 1998). This binding activity was shown to be due to a conserved cysteine residue in the HECT domain as mutation of this cysteine to alanine abolished ubiquitin binding activity. This section of the work enabled development of expertise in optimisation and analysis of *in vitro* translated and expressed recombinant protein.

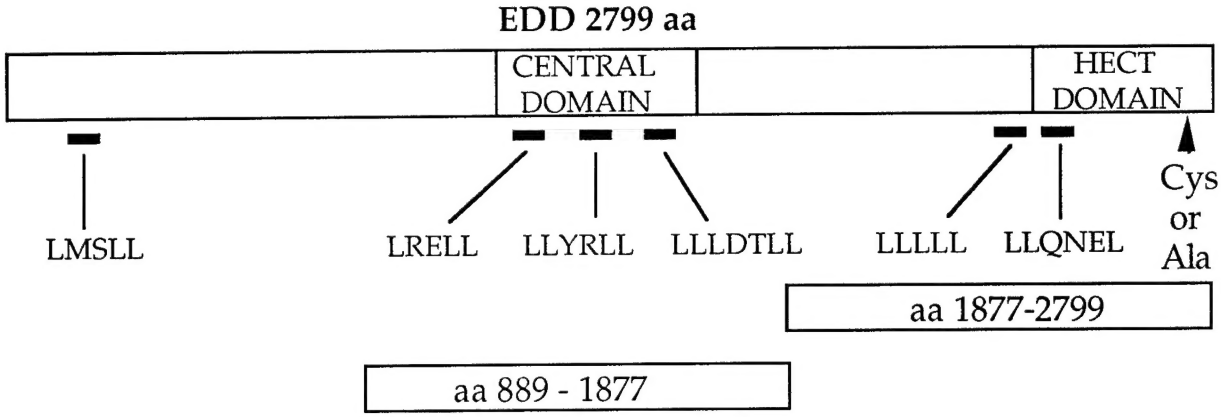


Figure 1. Schematic representation of the EDD protein sequence and overlapping deletion mutants.

Shown are the two domains highly conserved between EDD and HYD (CENTRAL and HECT) and the conserved cysteine residue (Cys) which has been mutated to alanine (Ala) in the ubiquitination-defective mutant. Three LLXXXL motifs, two of which coincide with two of the five LXXLL motifs, are also indicated. (aa, amino acids).

TASK 2: To identify substrates for EDD-mediated ubiquitination (months 6-36).

Twelve months into the project, there has already been significant progress towards this objective and the PI has gained valuable expertise in protein biochemistry and yeast two-hybrid approaches for studying protein-protein interactions. EDD was immunoprecipitated from T-47D whole cell lysates (800-1000 μg protein). Immunoprecipitates were resolved by SDS-PAGE and blotted for: (1) cell cycle regulatory molecules: cyclin D1, cyclin D3, cyclin E, CDK2, CDK4, p21, p27, pRB, p107 and p130; (2) cell signalling molecules: EMS1, 14.3.3β and Raf1; and (3) transcriptional regulators: ER, PR, RNA polII, SRC1 and p300. So far no interactions with these proteins has been reproducibly detected by co-immunoprecipitation of endogenous proteins.

EDD was also immunoprecipitated from metabolically labelled T-47D cell lysates and the pattern of bands compared to adjacent lanes from reactions where immunoprecipitation was peptide blocked. From these studies at least 3 proteins were consistently coimmunoprecipitated with EDD, of approx molecular weights 30 kDa, 70 kDa and 130 kDa. This has now been scaled up for silver staining of EDD-associated proteins in gels. Protein bands have been excised for identification by mass spectrometry in collaboration with the Garvan Institute's Protein Analysis Facility.

The main strategy for isolation of EDD-interacting proteins is to use yeast two hybrid library screening. Two human cDNA libraries, both cloned into the pACT2 vector, are available for screening in our laboratory, one derived from human placental mRNA and another derived from breast carcinoma mRNA. Full length EDD cDNA has been cloned downstream of the GAL4 DNA-binding domain in the pAS2.1 vector and used to screen three million clones from the placental cDNA library for interacting proteins. The C2768A mutant form of EDD is used to prevent degradation of potential interacting substrates. Two potentially interacting proteins, the human homologue of yeast SRP-1 (NPI-1 or karyopherin α) and calcium-integrin binding protein (CIB), were identified as proteins which enable reconstitution of GAL4 promoter-binding activity, thus driving *HIS3* and *LacZ* reporter gene expression in the yeast strain Y190. NPI-1 is a protein essential for nuclear transport and may regulate passage of EDD between the nucleus and cytoplasm, while CIB appears to be involved in cell signalling and also interacts with DNA-dependent protein kinase, suggesting a second, nuclear role. Antibodies and cDNAs have been obtained to confirm these interactions. GST pulldown confirmed interaction between EDD and NPI-1 and CIB. NPI-1 interacts most strongly with the amino two-thirds of EDD, which contains two bipartite nuclear localization signals while CIB binds to the carboxyl third of EDD, possibly through an Arg-Gly-Asp cell attachment sequence. EDD also co-immunoprecipitates with NPI-1 from T-47D cells. Studies are now underway to determine whether either of these proteins are ubiquitinated by EDD.

The use of the large full length EDD bait protein (~300 kDa) in this system may be problematic due to inefficient expression, explaining the low number of positives detected in the screen so two other baits have been constructed in the pAS2.1 vector. These consist of amino acids 889-1877 and 1877-2799 and span the two regions most highly conserved between EDD and HYD,

namely a 400 amino acid central domain and the HECT domain (with C2768A mutation), respectively (Fig.1). Expression of these constructs in yeast has been confirmed and they will be used to screen the available cDNA libraries for interacting proteins. Screening two million placenta library clones with the central domain bait has resulted in the isolation of a cDNA encoding a subunit of RNA pol II. This association may be important in light of the possible role of EDD as a transcriptional regulator (see below). In addition, five clones potentially encoding novel proteins were isolated.

The yeast two-hybrid system was also used to test for possible interactions with other proteins involved in the transcriptional complex such as the steroid receptor co-activators SRC-1, RAC3 and GRIP-1, TFIIβ, SKIP, and the nuclear receptors, PR, ER, VDR and RXRα in the presence or absence of ligand. EDD does not appear to interact directly with these proteins.

TASK 3: To determine the effects of EDD under- and over-expression on cultured breast cancer cells (months 6-24).

Wild type and mutant EDD were cloned in the inducible mammalian expression vectors pTRE and pΔMT and these vectors were used to transfect T-47D cells. No inducible expression of EDD was detected with either of these systems so as an alternative a human embryonic kidney cell line, HEK-293, previously shown in our laboratory to be amenable to protein expression studies, was transfected with EDD constructs in the pRcCMV constitutive vector. This has led to the establishment of several stable cell lines overexpressing either EDD or ubiquitination defective EDD, as well as empty vector transfected lines as controls. These lines have proven a valuable resource for characterising the EDD protein and for co-immunoprecipitation studies. We have preliminary data that EDD may affect the growth of the HEK-293 cells and are using these lines to confirm these observations. Obviously the studies on PR response in the absence of EDD are of high priority and we are working on antisense methods to regulate the levels of EDD in T-47D cells.

The HEK-293 expression system has enabled us to discover an additional and unexpected role for EDD: its apparent ability to function as a steroid receptor transcriptional co-regulator. The PR, like other members of the nuclear steroid receptor family, acts as a ligand-dependent transcriptional regulator by binding to specific promoter sequences and recruiting a variety of co-regulator proteins to its ligand-dependent C-terminal activation domain (AF-2). The proteins in the resulting transcriptional complex include: the steroid receptor co-activator family (SRC-1, GRIP-1, RAC3); TIF-1; RIP140; the co-integrator proteins CBP and p300; and p/CAF. These serve several functions: interacting with each other; controlling the assembly of the RNA polymerase II-containing pre-initiation complex and linking the activated receptor to it; and stimulating target gene chromatin remodelling through their intrinsic histone acetyl transferase activity. There is now considerable interest in the role that these co-regulators as well as co-repressor molecules (eg NCoR) might have in determining the gene- and cell type-specific actions of steroid hormones, and in hormone-dependent cancers, particularly in oncogenesis and in hormone resistance. AIB1 (SRC-3), for example, is commonly overexpressed and amplified in breast cancers, with higher expression in ER-positive breast cancer cell lines. Recently, the HECT-domain proteins yeast Rsp5 and its human homolog hRPF1 and E6-AP have been identified as having co-activator activity, potentiating progestin-dependent transcriptional activation of the PR. The mechanism for this activity is unknown. A search of the EDD sequence (see Fig. 1) reveals the presence of one N-terminal, one C-terminal and three centrally located LXXLL domains, which are involved in nuclear receptor interaction by some transcriptional co-activators, as well as two leucine-rich LLXXXL motifs that may function as SRC-family co-activator binding sites in nuclear receptor AF-2 domains. The centrally located motifs lie in a region of high homology to HYD. The presence of these motifs and the predominant nuclear localisation of EDD suggested to us that EDD might also possess co-activator activity and our preliminary experiments have confirmed this. Progestin response element (PRE)-containing reporter constructs were transiently transfected into HEK 293 cells, together with EDD and PR. EDD significantly enhanced the maximal ability of the synthetic progestin ORG2058 to stimulate PR reporter activity by up to 6-7 fold. No induction was observed with EDD alone, confirming a specific co-activating effect on the transfected PR. The ability of RU486 to act as a PR antagonist does not appear to be reversed by EDD, as is also the case for SRC-1, implying that EDD is unable to interact with the PR directly or indirectly when an antagonist occupies the receptor's ligand binding pocket, distinguishing it from basal transcription factors such as TATA Binding Protein. Because EDD may be a physiologically important regulator of the activity of PR and perhaps other transcription factors of direct relevance to hormone-dependent breast cancer, further functional analysis of this molecule and determination of its potential role in cancer is critical and a major research priority. Therefore the available overexpression system of

HEK-293 cells has been exploited to study the effects of EDD on PR activity as this may give us insights into the role of EDD in hormone dependent breast cancers. We are seeking funding to follow up this exciting preliminary data as part of an IDEAS grant.

TASK 4: To determine the expression of EDD in breast tumor specimens (months 6-18).

Initial attempts to perform immunohistochemistry on paraffin embedded cell blocks were not successful due to a high level of background staining by the antisera. The antisera has since been affinity purified and collaborations have been established with a local laboratory to test these new antibodies. When optimised, this technique will be applied to breast tumor specimens.

EDD mRNA was quantitated in 16 breast cancer cell lines and 2 normal breast cell lines by Northern blot analysis. Interestingly, the lowest levels were displayed by normal breast cell lines and EDD protein could not be detected in these cell lines. Breast tumor cell lines displayed a range of mRNA expression levels and all had detectable levels of EDD protein. While this may seem unexpected for a tumor suppressor protein such proteins are often normally present at low levels and upregulation can be a marker of mutation (eg p53 tumor suppressor protein). Further analysis of some cell lines by FISH showed amplification of the chromosomal region containing EDD in 2 of these lines.

Although it had been planned to quantitate EDD gene expression in breast tumor samples, given the possible tumor suppressor role of EDD, studies to determine the frequency of loss of heterozygosity or allelic imbalance at the EDD locus were given first priority. Given the availability to us of matched normal and tumor samples from a range of other cancers, these samples were used initially. We used microsatellite allelotyping of DNA extracted from tumors and matching normal tissues to determine whether chromosomal aberrations such as loss of heterozygosity (LOH) at the EDD locus on 8q22.3 (8) are common. Using a series of microsatellites mapped to the narrow region 8q22.2-8q23.3 including CEDD, a microsatellite within 100kb of EDD, we demonstrated that allelic imbalance occurred at high frequency, particularly involving CEDD, in ovarian cancers (notably in the serous subtype), hepatocellular carcinomas, squamous cell carcinomas of the tongue and metastatic melanomas (Table 1). In contrast, benign and borderline ovarian tumors had a very low frequency of allelic imbalance. These results are exciting as they are consistent with the presence of a tumor suppressor gene at or very near the EDD locus and suggest that the EDD gene may have a common role in the progression of several human cancers. This approach will be extended to breast cancer where to date the only relevant data we have indicate that 2 of 16 breast cancer cell lines have apparent LOH at the EDD locus as judged by homozygosity for all six microsatellite markers. Collaborative studies have already commenced to determine EDD gene expression in breast tumor specimens by *in situ* hybridisation.

Table 1: Frequency of allelic imbalance in human cancers.

The number of tumors showing allelic imbalance as a proportion of the number of tumors informative for each microsatellite marker is displayed for four tumor types. Totals are across all four cancer types. Markers span the region 8q22.2-23.3 and the microsatellite CEDD is an EDD-linked microsatellite located within 100kb of the EDD gene (8q22.3).

Microsatellite Marker	Ovarian	Liver	Melanoma	Tongue	Total	%
D8S257	17/40	2/15	4/13	0/7	23/75	31
CEDD	22/46	7/14	2/11	2/4	33/75	44
D8S326	22/55	8/14	5/16	4/7	39/92	42
D8S300	12/31	6/16	3/15	N/A	21/62	34
D8S545	13/45	5/16	5/13	0/4	23/78	30
D8S85	13/39	4/15	4/15	0/9	21/78	27

APPENDICES

1. Key Research Accomplishments

- Demonstration of the ability of EDD to reversibly bind ubiquitin through a conserved cysteine residue
- Generation of antisera for immunoprecipitation and Western blotting of the EDD protein
- Characterisation of the EDD protein by cloning and expression of full length recombinant protein and publication of these results
- Eighteen candidate proteins were tested for interaction with EDD by immunoprecipitation
- Nine proteins involved in transcriptional complex formation were tested for interaction with EDD using the yeast two-hybrid system.
- EDD constructs were made and used to screen a placental cDNA library in the yeast two-hybrid system. Eight interacting proteins, including five novel proteins, were identified and two of these have already been confirmed as EDD-interacting proteins.
- The Tet-Off system was tested for inducible EDD expression in T-47D cells but found to be unsuccessful.
- An alternative expression system has been established in HEK-293 cells and an initiative undertaken to produce antisense EDD constructs for downregulation of the EDD protein in T-47D cells.
- Discovery of an exciting new role for EDD in transcriptional activation by progesterone receptor.
- Data on the expression of EDD in 16 breast cancer and 2 normal breast cell lines has been gathered.
- Evidence of a role for EDD abnormalities in the evolution and/or development of several tumor types

2. Reportable Outcomes

Manuscripts

Callaghan MJ, Russell AJ, Woollatt E, Sutherland GR, Sutherland RL and Watts CKW (1998). Identification of a human HECT family protein with homology to the *Drosophila* tumor suppressor gene *hyperplastic discs*. **Oncogene 17, 3479-3491.**

Poster presentations

1. Identification of a progestin-regulated gene with homology to the *Drosophila* tumour suppressor gene *hyperplastic discs*
   Amanda J. Russell, Michelle J. Callaghan, Robert L. Sutherland and Colin K. W. Watts.
   8th St Vincent's Campus Research Symposium, September 11th, 1998, St Vincent's Hospital, Sydney, NSW, Australia.

2. Identification of a progestin-regulated gene with homology to the *Drosophila* tumour suppressor gene *hyperplastic discs*
   Michelle J. Henderson , Jennifer Clancy, Amanda J. Russell, Robert L. Sutherland and Colin K. W. Watts.
   11th Lorne Cancer Conference, February 11th-14th, 1999, Lorne, Victoria, Australia.

Oral presentations

1. Identification of a progestin-regulated gene with homology to the *Drosophila* tumour suppressor gene *hyperplastic discs*
   Michelle J. Callaghan, Amanda J. Russell, Robert L. Sutherland and Colin K. W. Watts.
   Invited Speaker, Endocrine Society of Australia Annual Scientific Meeting, Satellite Symposium on Steroid Hormone Receptors and Cancer, August 26th-27th, 1998, Perth, WA, Australia.

2. A novel progestin-regulated gene isolated from breast cancer cells.
Michelle J. Callaghan, Amanda J. Russell, Robert L. Sutherland and Colin K. W. Watts.
Nuclear Receptors Symposium, September 11th, 1998, Garvan Institute of Medical Research, Sydney, NSW, Australia.

3. Identification of a progestin-regulated gene with homology to the *Drosophila* tumour suppressor gene *hyperplastic discs*
Colin K. W. Watts, Michelle J. Callaghan, Amanda J. Russell and Robert L. Sutherland.
Invited Speaker, New Horizons in Breast Cancer Research Conference, November 16th-18th, 1998, University of Melbourne, Victoria, Australia.

Development of cell lines
Clonal lines of HEK-293 cells stably transfected with pRcCMV vector (10 lines), or pRcCMV vector containing EDD (3 lines) or EDD ubiquitination defective mutant (4 lines) have been established.

Funding applications (requested)
The progress made so far in this project has enabled the preparation of several new grant applications.
1. Title: Investigation of a putative tumour suppressor gene with ubiquitin-protein ligase and co-activator functions
Agency: The Leo & Jenny Leukemia and Cancer Foundation of Australia
PI: Dr. Michelle Henderson

2. Title: Functional characterisation of a putative tumour suppressor gene and its role in human cancer
Agency: New South Wales Cancer Council
PIs: Dr. Michelle Henderson, Dr. Colin Watts

3. Title: Functional characterisation of a novel progesterone receptor co-activator and its role in breast cancer.
Agency: US Army Breast Cancer Research Program IDEAS Awards
PI: Dr. Colin Watts, AI: Dr. Michelle Henderson

Research Opportunities
The progress made in this project resulted in the participation of the PI in a national conference (Endocrine Society of Australia) as an invited speaker. In addition, important collaborative ties have been made possible.

Oncogene (1998) 17, 3479–3491

http://www.stockton-press.co.uk/onc

Identification of a human HECT family protein with homology to the *Drosophila* tumor suppressor gene *hyperplastic discs*

Michelle J Callaghan[1], Amanda J Russell[1], Erica Woollatt[2], Grant R Sutherland[2], Robert L Sutherland[1] and Colin KW Watts[1]

[1]*Cancer Research Program, Garvan Institute of Medical Research, St Vincent's Hospital, Sydney, NSW 2010;* [2]*Centre for Medical Genetics, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, Adelaide, SA 5006, Australia*

Use of the differential display technique to isolate progestin-regulated genes in T-47D human breast cancer cells led to identification of a novel gene, EDD. The cDNA sequence contains a 2799 amino acid open reading frame sharing 40% identity with the predicted 2894 amino acid product of the *Drosophila melanogaster* tumor suppressor gene *hyperplastic discs*, while the carboxy-terminal 889 amino acids show 96% identity to a rat 100 kDa HECT domain protein. EDD mRNA was progestin-induced in T-47D cells and was highly abundant in testes and expressed at moderately high levels in other tissues, suggesting a broad role for EDD. Anti-EDD antibodies immunoprecipitated an approximately 300 kDa protein from T-47D cell lysates. HECT family proteins function as E3 ubiquitin-protein ligases, targeting specific proteins for ubiquitin-mediated proteolysis. EDD is likely to function as an E3 as *in vitro* translated protein bound ubiquitin reversibly through a conserved HECT domain cysteine residue. EDD was localized by FISH to chromosome 8q22, a locus disrupted in a variety of cancers. Given the homology between EDD and the *hyperplastic discs* protein, which is required for control of imaginal disc growth in *Drosophila*, EDD potentially has a role in regulation of cell proliferation or differentiation.

Keywords: ubiquitination; tumor suppressor; progestin; breast cancer

Introduction

The control of cell proliferation and differentiation in the normal breast and in breast cancer involves complex actions and interactions of steroid hormones (in particular estrogen and progesterone), peptide hormones and growth factors (Clarke and Sutherland, 1990; Dickson and Lippman, 1995). How these agents act at critical control points within the cell cycle to influence progression through the cycle or exit to enter a pathway of differentiation is only partially understood (Musgrove *et al.*, 1991; Prall *et al.*, 1997; Musgrove *et al.*, 1998). Progestins are responsible for mammary gland lobuloalveolar development during pregnancy (Lydon *et al.*, 1995), although there is evidence for a more predominant role for estrogens than progestins in stimulating epithelial cell proliferation in the normal premenopausal breast (Laidlaw *et al.*, 1995; Sutherland *et al.*, 1998) Progestins both stimulate and inhibit breast cancer epithelial cell proliferation *in vitro* but the predominant effect is growth inhibition probably *via* induction of differentiation (Musgrove *et al.*, 1991; Groshong *et al.*, 1997; Musgrove *et al.*, 1998; Sutherland *et al.*, 1998)

Progestin action is mediated primarily through the progesterone receptor (PR), which acts as a transcriptional transactivator for a largely undefined set of progestin-responsive genes which may in turn transcriptionally or post-transcriptionally influence additional genes or gene products. Only a limited number of genes have been implicated in progestin action on cell proliferation. Studies from this laboratory have identified c-*myc* and cyclin D1 as major downstream targets of progestin-stimulated cell cycle progression in human breast cancer cells (Musgrove *et al.*, 1991; 1993) while the delayed growth inhibitory effects of progestins involve decreases in cyclin D1 and E gene expression (Groshong *et al.*, 1997; Musgrove *et al.*, 1998). While progestin effects on c-*myc* gene expression are rapid and occur within minutes, effects on cyclin expression begin several hours later, pointing to the presence of undefined earlier events.

Since progestin action is complex and is likely to involve multiple genes, many of which are currently unknown, the differential display RT–PCR technique (DD–PCR) (Liang and Pardee, 1992) was adopted to identify target genes in cultured human breast cancer cells. We have previously demonstrated the utility of this approach by the cloning of PRG1, a gene having significant homology with isoforms of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (Hamilton *et al.*, 1997). Using the same technique a novel progestin-regulated gene, EDD, has been identified. Based on the sequence similarity reported here, EDD appears to be a human homolog of the *Drosophila* tumor suppressor gene *hyperplastic discs* (*hyd*) (Mansfield *et al.*, 1994). Although the function of the HYD protein is unknown, significant homology exists between its carboxyl terminus and those of human E6-AP and a number of proteins identified through database searches (Huibregtse *et al.*, 1995). These HECT domain (*h*omology to *E*6-AP *c*arboxyl *t*erminus) family proteins function as ubiquitin-protein ligases (E3 enzymes) (Huibregtse *et al.*, 1995; 1997; Hatakeyama *et al.*, 1997), playing a role in the ubiquitination cascade that targets specific substrate proteins for proteolysis. Only a few mammalian E3 genes have been cloned to date, namely E6-AP

Correspondence: CKW Watts
Received 29 May 1998; revised 26 June 1998; accepted 26 June 1998

(Huibregtse *et al*., 1993; Scheffner *et al*., 1993), UreB1 (Gu *et al*., 1994), *Nedd4* (Kumar *et al*., 1997a) and Itch (Perry *et al*., 1998).

In the present study the cDNA cloning and initial characterization of a new human E3, EDD (*E*3 identified by *D*ifferential *D*isplay), is reported. The EDD protein has a carboxy-terminal HECT domain containing a cysteine residue that covalently binds ubiquitin. As this amino acid is conserved in all known HECT domain-containing E3 enzymes, and is involved in the transfer of ubiquitin, it is proposed that EDD is a novel human ubiquitin-protein ligase.

Results

Isolation and Northern blot analysis of a progestin regulated cDNA

The differential display technique was used to identify mRNAs in T-47D human breast cancer cells with altered levels of expression in response to treatment with the synthetic progestin ORG 2058 for 3 h. When the anchored primer, 5′ACGACTCACTATAG-$GGCT_{12}AC$ was used in conjunction with the arbitrary primer, 5′ACAATTTCACACAGGAGCTAGCAGAC, a cDNA fragment of approximately 850 bp that was more abundant in treated samples than in control samples was identified and designated EDD (Figure 1a). Northern analysis of total cellular RNA from T-47D cells showed that transcription was required for the observed ORG 2058 induction of EDD mRNA levels as this was blocked in the presence of actinomycin D (Figure 1b). Induction was also prevented by cycloheximide, suggesting that EDD is not directly transcriptionally regulated by progestin acting *via* the PR (Figure 1c).

The tissue specificity of EDD gene expression was investigated by hybridizing Northern blots of $polyA^+$ RNA isolated from human tissues with the EDD cDNA fragment. A single transcript of 9.5 kb was detected in a variety of tissues (Figure 2a) with the highest expression in testis, heart, placenta and skeletal muscle. Hybridization to a more quantitatively loaded RNA dot blot (Figure 2b) confirmed that EDD is expressed at varying levels in all tissues examined and that the mRNA was most abundant in testis and expressed at high levels in brain, pituitary and kidney.

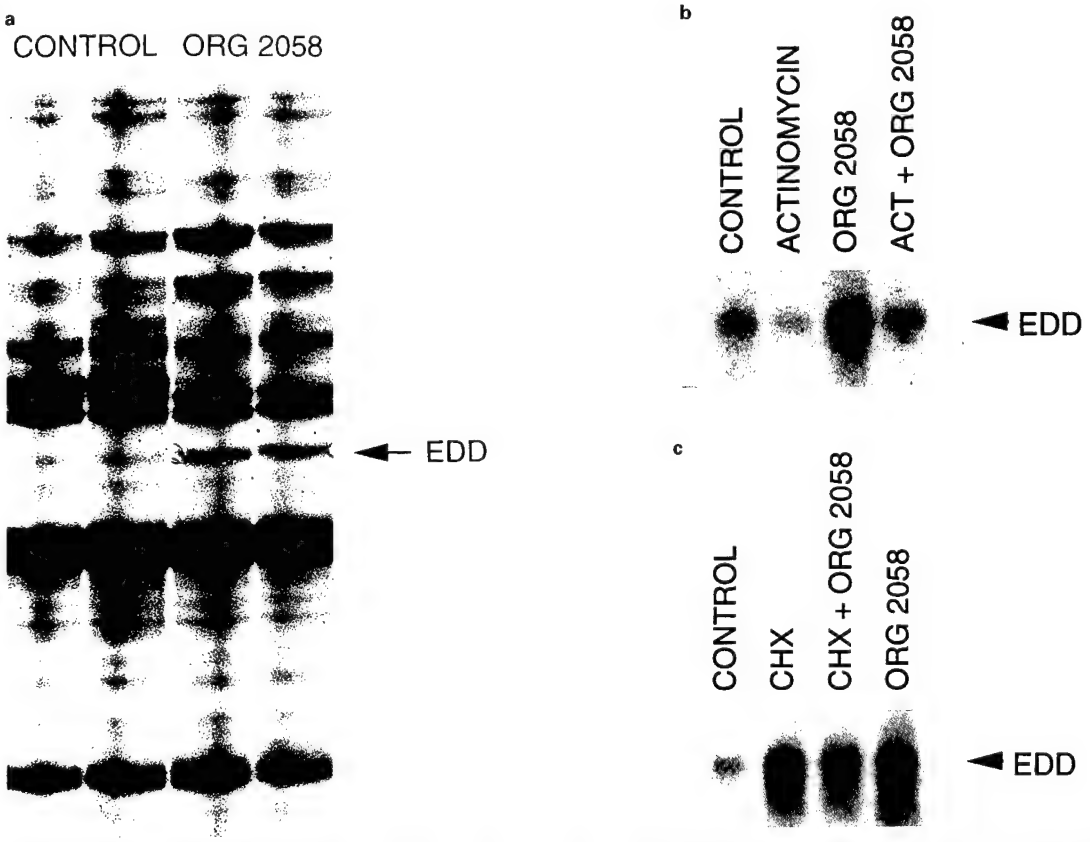


Figure 1 Identification of a differentially expressed cDNA in T-47D cells treated with the synthetic progestin ORG 2058. (**a**) Identification of EDD by differential display. Total RNA obtained from T-47D cells treated with ORG 2058 or vehicle control (ethanol) for 3 h was used as a template for differential display PCR reactions. The PCR products were separated on a 4.5% polyacrylamide denaturing gel and visualized by autoradiography. The arrow indicates the EDD DD-PCR product (DD5-1; see Figure 3a) which is present at a higher level in the progestin treated (ORG 2058) compared with control lane. (**b**) Confirmation of the progestin induction of EDD by Northern blot analysis. T-47D cells proliferating in medium supplemented with 5% charcoal-treated FCS were treated with 10 nM ORG 2058 or ethanol vehicle (CONTROL) in the presence or absence of actinomycin D (ACT) and after 3 h total RNA was harvested for Northern analysis. The Northern blot was probed with the EDD clone P19. (**c**) Effect of cycloheximide on progestin induction of EDD mRNA. T-47D cells proliferating in medium supplemented with 5% charcoal-treated FCS were treated with ORG 2058 (10 nM), cycloheximide (CHX, 20 μg/ml), ORG 2058 and CHX simultaneously or ethanol vehicle and harvested for total RNA at 1 h. The Northern blot was probed with the EDD DD-PCR fragment DD5-1

Significant levels of expression were also observed in placenta, uterus, prostate, stomach, fetal lung and various brain tissues. EDD mRNA was also expressed in a range of breast cancer cell lines, not all of which are progestin-responsive (not shown).

Cloning of the full length EDD cDNA

The original DD5-1 fragment isolated by DD PCR was 850 bp in length and is shown schematically in Figure 3a. The DNA sequence of this fragment had no homology to sequences of any known human genes. To obtain the complete coding sequence from which EDD was derived a combination of 5′RACE and screening of human heart and placenta cDNA libraries was used. This resulted in a series of overlapping clones covering 8.5 kb of sequence (Figure 3a; Genbank Accession AF006010). Analysis of the sequence revealed an open reading frame of 2799 amino acids (Figure 3b). The EDD sequence was divided into overlapping 1800 bp segments and used in Blastx searches of the GenBank database. The only homology to a human sequence of known function was to polyA binding protein across 50 amino acids (50%, Figure 3b) although the similarities among mammalian polyA binding proteins in this stretch are usually in the vicinity of 100%.

The DNA sequence of EDD showed significant similarity to two sequences in the database. Both of

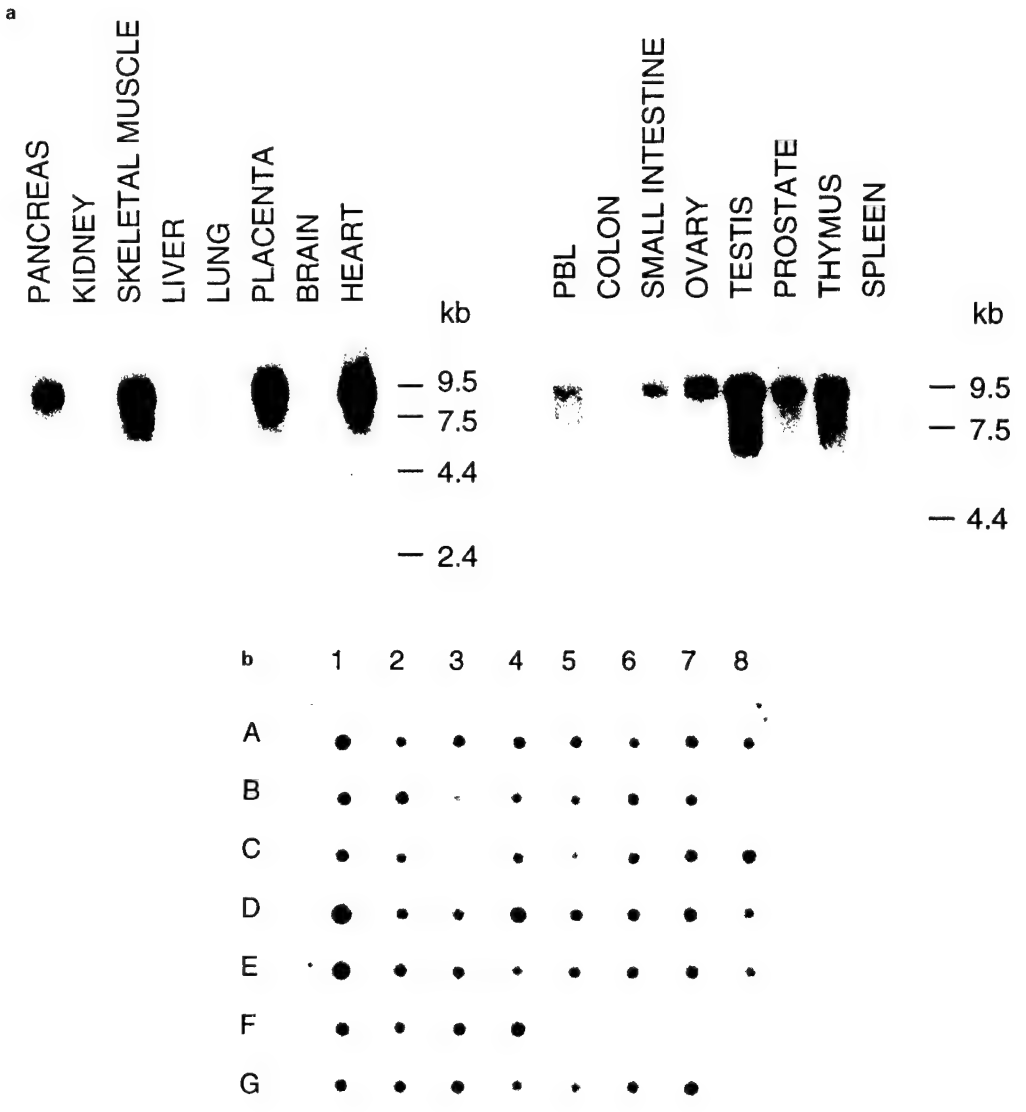
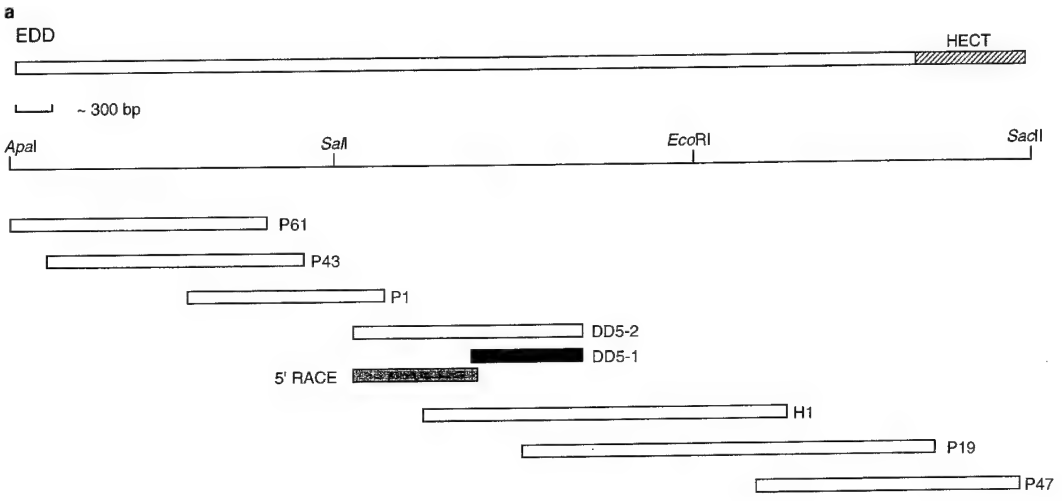


Figure 2 Expression of EDD mRNA in human tissues. (**a**) Northern blot analysis of poly A^+ RNA from human tissues. The blot was hybridized with the P19 cDNA clone of EDD. Molecular sizes of markers are indicated. PBL, peripheral blood leukocytes. (**b**) Dot blot analysis of poly A^+ RNA from human tissues. The blot was hybridized with the P19 cDNA clone of EDD. Row A: 1, whole brain; 2, amygdala; 3, caudate nucleus; 4, cerebellum; 5, cerebral cortex; 6, frontal lobe; 7, hippocampus; 8, medulla oblongata; Row B: 1, occipital lobe; 2, putamen; 3, substantia nigra; 4, temporal lobe; 5, thalamus; 6, sub-thalamic nucleus; 7, spinal cord; Row C: 1, heart; 2, aorta; 3, skeletal muscle; 4, colon, 5, bladder; 6, uterus; 7, prostate; 8, stomach; Row D: 1, testis; 2, ovary; 3, pancreas; 4, pituitary gland; 5, adrenal gland; 6, thyroid gland; 7, salivary gland; 8, mammary gland; Row E: 1, kidney; 2, liver; 3, small intestine; 4, spleen; 5, thymus; 6, peripheral leukocyte; 7, lymph node; 8, bone marrow; Row F: 1, appendix; 2, lung; 3, trachea; 4, placenta; Row G: 1, fetal brain; 2, fetal heart; 3, fetal kidney; 4, fetal liver; 5, fetal spleen; 6, fetal thymus; 7, fetal lung

a



b

```
1    MTSIHFVVHP LPGTEDQLND RLREVSEKLN KYNLNSHPPL NVLEQATIKQ CVVGPNHAAF
61   LLEDGRVCRI GFSVQPDRLE LGKPDNNDGS KLNSNSGAGR TSRPGRTSDS PWFLSGSETL
121  GRLAGNTLGS RWSSGVGGSG GGSSGRSSAG ARDSRRQTRV IRTGRDRGSG LLGSQPQPVI
181  PASVIPEELI SQAQVVLQGK SRSVIIRELQ RTNLDVNLAV NNLLSRDDED GDDGDDTASE
241  SYLAGEDLMS LLDADIHSAH PSVIIDADAM FSEDISYFGY PSFRRSSLSR LGSSRVLLLP
301  LERDSELLRE RESVLRLRER RWLDGASFDN ERGSTSKEGE PNLDKKNTPV QSPVSLGEDL
361  QWWPDKDGTK FICIGALYSE LLAVSSKGEL YQWKWSESEP YRNAQNPSLH HPRATFLGLT
421  NEKIVLLSAN SIRATVATEN NKVATWVDET LSSVASKLEH TAQTYSELQG ERIVSLHCCA
481  LYTCAQLENS LYWWGVVPFS QRKKMLEKAR AKNKKPKSSA GISSMPNITV GTQVCLRNNP
541  LYHAGAVAFS ISAGIPKVGV LMESVWNMND SCRFQLRSPE SLKNMEKASK TTEAKPESKQ
601  EPVKTEMGPP PSPASTCSDA SSIASSASMP YKRRRSTPAP KEEEKVNEEQ WSLREVVFVE
661  DVKNVPVGKV LKVDGAYVAV KFPGTSSNTN CQNSSGPDAD PSSLLQDCRL LRIDELQVVK
721  TGGTPKVPDC FQRTPKKLCI PEKTEILAVN VDSKGVHAVL KTGNWVRYCI FDLATGKAEQ
781  ENNFPTSSIA FLGQNERNVA IFTAGQESPI ILRDGNGTIY PMAKDCMGGI RDPDWLDLPP
841  ISSLGMGVHS LINLPANSTI KKKAAVIIMA VEKQTLMQHI LRCDYEACRQ YLMNLEQAVV
901  LEQNLQMLQT FISHRCDGNR NILHACVSVC FPTSNKETKE EEEAERSERN TFAERLSAVE
961  AIANAISVVS SNGPGNRAGS SSSRSLRLRE MMRRSLRAAG LGRHEAGASS SDHQDPVSPP
1021 IAPPSWVPDP PAMDPDGDID FILAPAVGSL TTAATGTGQG PSTSTIPGPS TEPSVVESKD
1081 RKANAHFILK LLCDSVVLQP YLRELLSAKD ARGMTPFMSA VSGRAYPAAI TILETAQKIA
1141 KAEISSSEKE EDVFMGMVCP SGTNPDDSPL YVLCCNDTCS FTWTGAEHIN QDIFECRTCG
1201 LLESLCCCTE CARVCHKGHD CKLKRTSPTA YCDCWEKCKC KTLIAGQKSA RLDLLYRLLT
1261 ATNLVTLPNS RGEHLLLFLV QTVARQTVEH CQYRPPRIRE DRNRKTASPE DSDMPDHDLE
1321 PPRFAQLALE RVLQDWNALK SMIMFGSQEN KDPLSASSRI GHLLPEEQVY LNQQSGTIRL
1381 DCFTHCLIVK CTADILLLDT LLGTLVKELQ NKYTPGRREE AIAVTMRFLR SVARVFVILS
1441 VEMASSKKKN NFIPQPIGKC KRVFQALLPY AVEELCNVAE SLIVPVRMGI ARPTAPFTLA
1501 STSIDAMQGS EELFSVEPLP PRPSSDQSSS SSQSQSSYII RNPQQRRISQ SQPVRGRDEE
1561 QDDIVSADVE EVEVVEGVAG EEDHHDEQEE HGEENAEAEG QHDEHDEDGS DMELDLLAAA
1621 ETESDSESNH SNQDNASGRR SVVTAATAGS EAGASSVPAF FSEDDSQSND SSDSDSSSSQ
1681 SDDIEQETFM LDEPLERTTN SSHANGAAQA PRSMQWAVRN TQHQRAASTA PSSTSTPAAS
1741 SAGLIYIDPS NLRRSGTIST SAAAAAAALE ASNASSYLTS ASSLARAYSI VIRQISDLMG
1801 LIPKYNHLVY SQIPAAVKLT YQDAVNLQNY VEEKLIPTWN WMVSIMDSTE AQLRYGSALA
1861 SAGDPGHPNH PLHASQNSAR RERMTAREEA SLRTLEGRRR ATLLSARQGM MSARGDFLNY
1921 ALSLMRSHND EHSDVLPVLD VCSLKHVAYV FQALIYWIKA MNQQTTLDTP QLERKRTREL
1981 LELGIDNEDS EHENDDDTNQ SATLNDKDDD SLPAETGQNH PFFRRSDSMT FLGCIPPNPF
2041 EVPLAEAIPL ADQPHLLQPN ARKEDLFGRP SQGLYSSSAS SGKCLMEVTV DRNCLEVLPT
2101 KMSYAANLKN VMNMQNRQKK EGEEQPVLPE ETESSKPGPS AHDLAAQLKS SLLAEIGLTE
2161 SEGPPLTSFR PQCSFMGMVI SHDMLLGRWR LSLELFGRVF MEDVGAEPGS ILTELGGFEV
2221 KESKFRREME KLRNQQSRDL SLEVDRDRDL LIQQTMRQLN NHFGRRCATT PMAVHRVKVT
2281 FKDEPGEGSG VARSFYTAIA QAFLSNEKLP NLECIQNANK GTHTSLMQRL RNRGERDRER
2341 ERREREMRRSS GLRAGSRRDR DRDFRRQLSI DTRPFRPASE GNPSDDPEPL PAHRQALGER
2401 LYPRVQAMQP AFASKITGML LELSPAQLLL LLASEDSLRA RVDEAMELII AHGRENGADS
2461 ILDLGLVDSS EKVQQENRKR HGSSRSVVDM DLDDTDDGDD NAPLFYQPGK RGFYTPRPGK
2521 NTEARLNCFR NIGRILGLCL LQNELCPITL NRHVIKVLLG RKVNWHDFAF FDPVMYESLR
2581 QLILASQSSD ADAVFSAMDL AFAIDLCKEE GGGQVELIPN GVNIPVTPQN VYEYVRKYAE
2641 HRMLVVAEQP LHAMRKGLLD VLPKNSLEDL TAEDFRLLVN GCGEVNVQML ISFTSFNDES
2701 GENAEKLLQF KRWFWSIVEK MSMTERQDLV YFWTSSPSLP ASEEGFQPMP SITIRPPDDQ
2761 HLPTANTCTS RLYVPLYSSK QILKQKLLLA IKTKNFGFV
```

Figure 3 Cloning and predicted amino acid sequence of EDD. **(a)** A schematic representation of EDD structure with a restriction map for the EDD cDNA indicating the sites used for cloning the full-length EDD construct and the cDNA clones used to derive the EDD sequence shown beneath. The DD-PCR cDNA fragment identified by differential display was designated DD5-1 and a cDNA clone derived from the 5′ RACE product and the original DD-PCR product, DD5-2. All cDNA clones were isolated from a human placenta cDNA library with the exception of H1 which was isolated from a human heart cDNA library. **(b)** Predicted amino acid sequence of EDD. There are two regions with high homology (~60%) to HYD (a central sequence and a carboxyl sequence containing the HECT domain) and these and other highly conserved sequences are shown in bold type, while two putative nucelar localization signals are boxed. The HECT domain is underlined and in bold type and includes a conserved cysteine at residue 2768 (boxed). A region showing homology to polyA-binding proteins is italicized (aa 2395–2445) and the peptide sequence to which antiserum AbPEP1 was raised is underlined (aa 2469–2484). The numbers refer to positions of amino acids

these genes encode proteins belonging to the HECT family of ubiquitin-protein ligases, although their specificities are unknown. HECT proteins contain a conserved domain of approximately 300 amino acids that contains a cysteine residue able to bind ubiquitin via a thioester linkage. Nucleotides 5667 to 8502 of EDD were 88% identical to the rat 100 kDa protein cDNA sequence (Muller *et al.*, 1992), nucleotides 572–740 and 3498–3867 were 69% identical to two regions of the *Drosophila melanogaster hyperplastic discs* gene (*hyd*) and nucleotides 7560–8430 were 60% identical to a third region of *hyd* (Mansfield *et al.*, 1994). The putative initiation codon is surrounded by a consensus sequence for strong translational initiation (ACCATGA (Kozak, 1987)) and corresponds to a possible start codon of the Drosophila *hyd* gene (Mansfield *et al.*, 1994). The stop codon corresponds to that shared by the rat 100 kDa protein and *hyd* genes. Like EDD, both the *hyd* and rat 100 kDa protein genes have estimated mRNA transcript sizes of 9.5 kb (Muller *et al.*, 1992; Huibregtse *et al.*, 1995). The predicted EDD protein is identical to HYD at 40% of amino acid residues and similar at 64% of residues, while the carboxyl third of EDD is 96% identical and 98.5% similar to rat 100 kDa protein. The most highly conserved regions between HYD and EDD are designated by bold type in Figure 3b. The longest conserved regions between EDD and HYD are a central domain of approximately 400 amino acids (58% identity, 72% similarity) and the carboxyl 300 amino acids which include the HECT domain and conserved cysteine residue (64% identity, 80% similarity). These regions also show considerable similarity to a possible *C. elegans* homolog of HYD contained within two overlapping cosmids (Genbank Accession No. G2652994 and G2653017) and the alignments of these sequences are shown in Figure 4a and b. The HECT domain showed around 30% identity and 50% similarity with other HECT proteins including yeast RSP5 or PUB-1 and RAD26 (van Gool *et al.*, 1994; Huibregtse *et al.*, 1995; Nefsky and Beach, 1996), and the mammalian proteins UreB1 (Gu *et al.*, 1994), *Nedd-4* (Nagase *et al.*, 1995; Staub *et al.*, 1996; Hatakeyama *et al.*, 1997; Kumar *et al.*, 1997a) and E6-AP (Huibregtse *et al.*, 1993; Scheffner *et al.*, 1993; Hatakeyama *et al.*, 1997). Apart from two putative nuclear localization signals (Dingwall and Laskey, 1991), no other consensus functional domains were identified within the EDD sequence.

Chromosomal localization of the EDD gene

FISH was used to localize the gene for EDD. Eighteen metaphases from a normal male were examined for fluorescent signal. Seventeen of these metaphases showed signal on one or both chromatids of chromosome eight in the region q22. High resolution studies of eight metaphases showed signal at q22.3 (Figure 5). There was a total of four non-specific background dots observed in these 18 metaphases. A similar result was obtained from hybridization of the probe to 11 metaphases from a second normal male (data not shown). This localization was consistent with independent assignment of an EST corresponding to EDD (EST116344) using the radiation hybrid panel Genebridge 4.

Characterization of EDD protein

A rabbit antiserum (AbPEP1) was raised against an EDD-specific peptide matching a sequence towards the carboxyl terminus of the protein (underlined in Figure 3b). Strong cross-reactivity with EDD was demonstrated by Western blotting of a truncated (100 kDa) recombinant protein expressed in *Sf9* cells using a baculovirus system (Figure 6a). A second strongly immunoreactive band of approximately 200 kDa was also seen, but this appeared to be non-specific as antibody binding was not competed by the EDD peptide. The full length EDD cDNA was cloned into pBluescript and translated *in vitro* in a rabbit reticulocyte lysate system. The size of the major product was in agreement with the expected molecular mass of the protein as predicted from the amino acid sequence (~300 kDa, Figure 6b). The identity of the translated protein was confirmed by immunoprecipitation from either translation reactions or T-47D whole cell lysates with AbPEP1 (Figure 6b). Western blotting of whole cell lysates from T-47D cells using AbPEP1 detected two major bands, both abolished in the presence of competing peptide–a major species at approximately 230 kDa and a minor species of higher molecular mass (Figure 6c). This latter band corresponds in size to that of the *in vitro* translated protein and is immunoprecipitated by AbPEP1 (Figure 6c) and by two other EDD-specific peptide antibodies (not shown). However, the 230 kDa protein is not immunoprecipitated from cell lysates by these antibodies. As a single EDD mRNA transcript was detected on Northern blots, we at first hypothesized that the EDD protein may be processed to the 230 kDa form which is folded in a way that is not susceptible to immunoprecipitation in its native state. However, transient expression of full length EDD in HEK-293 cells resulted in a large increase in the 300 kDa species relative to endogenous levels, whereas levels of the 230 kDa protein remained unchanged (Figure 6d), strongly suggesting that the 230 kDa protein is not an EDD gene product. Western blotting of whole cell lysates from a number of normal breast and breast cancer epithelial cell lines showed that EDD protein was expressed in all immortalized and cancer cell lines but not in a normal breast cell line, 184 (Figure 7).

Identity of the rat gene product

The previously described rat cDNA that is highly homologous to the EDD gene reportedly gives rise to a 100 kDa protein, inferred from cDNA sequence data which showed several in-frame stop codons upstream of the putative initiation codon (Muller *et al.*, 1992), corresponding to amino acid residue 1910 of EDD. These stop codons were not present in the EDD cDNA. Furthermore, although we were able to confirm that an anti-HYD antibody detected an approximately 100 kDa protein in rat muscle lysates, this species was not detected by AbPEP1 even though the predicted sequences of human and rat proteins are identical at every residue of the peptide used to raise the AbPEP1 antibody. This led us to question whether the 100 kDa protein was the actual rat gene product.

a

```
          1102                                                    1151
EDD       LRELLSAKDA RGMTPFMSAV SGRAYPAAIT ILETAQKIAK AEISSSEKEE
HYD       LCHMLSTKDA QGQTPFMLSV SCRAYEAGII LLNTILML.. ....SEQ.DP
C.ELEGANS VKALLLTRDL NGMTPFQSAI NQRAYGAATS IWRALLQLGF QTMSEEQKLK

          1152                                                    1201
EDD       DVFMGMVCPS GTNPDDSPLY VLCCNDTCSF TWTGAEHINQ D:FECRTCGL
HYD       QLKEAMIFPN GSPADQSPLH VICYNDTCSF TWTGADHINQ N:FECKTCGL
C.ELEGANS YIFPSFNVDE SHNAD.SPLF ILCYNDVCSF TWTGEDHINQ D:YECKTCGL

          1202                                                    1249
EDD       LESLCCCTEC AGVCHKGHDW KLKRTSPTAY CDCWEK..CK CI:TLIAGQKS
HYD       TGSLCCCTEC ARVCHKGHDC KLKRTAPTAY CDCWEK..CK CI:ALIAGNL.
C.ELEGANS TGSLCCCSEC ALTCHRNHDC RLKRTSPTAY CDCWEKSSCN CI:ALIAGNEL

          1250                                                    1298
EDD       ARLDLLYRLL TATNLVTLPN SRGEHLLLFL VQTVARQTVE H:QYR.PPRI
HYD       TKLXLLCKLV SCTDLVTKFN SKGESILLFL IQTVGRQIVE Q:QYRFSVRV
C.ELEGANS LREHLLSELL KYTRLFEYPN ARNEHIMLFL ARTVIRQTRE Q:...YSQRR

          1299                                                    1333
EDD       RE..... .....DRNR KTASPE.DSD MPDHDLEPPR F:QLALERVL
HYD       RNVSTAATGA TGNNSVISNR KTSAAEIDND MPDHDLEPPK F:RKALERLL
C.ELEGANS RSRYTAM... ..... ...QTSQSSG TPEHNLNPPK F:KTALEQCC

          1334                                                    1382
EDD       QDWNALKSMI MFGSQENKDP LSASSRIGHL LPEE.QVYLN Q:SGTIRLDC
HYD       IDWNAVRSMI MSGAERGDVP NPAGSASENS NSEGFNMFIQ T:HGSTLLDK
C.ELEGANS KNANVVMNSI RCGINKFGQM ELATPG.... .....QLHFV Q:STSTHLDK

          1383                                                    1431
EDD       FTHCLIVKCT ADILLLDTLL GTLVKELQNK YTPGRREEAI A:'.TMRFLRS
HYD       FTHSLIVKCT SD..HLDTLL LTLVRELQNA SVSNRSKEAE E:'.VRRFVRS
C.ELEGANS LTFALSNKIL KDYSSL..FV GTLRNLLSAP IRDPRVDQDI E:WVGRFVRS

          1432                                                    1475
EDD       VARVFVILSV EMASSKKKNN FIPQP...IG KCKRV...FQ A:LPYAVEEL
HYD       VARAFVIFNL QKQPNXQKQK SHSSCNKYVQ SCVKV...FQ T:HKISIEEL
C.ELEGANS LVRVLTVFHA LVISKKESSK LSEKSASFVA MITRIMQFLQ Q:PTFAVVQL

          1476
EDD       CNVAESLIVP VRMGIARP
HYD       CEVSEALIAP VRLGVVRP
C.ELEGANS AIAADAIFEP VRDGMLRP
```

b

```
          2474                                                    2523
EDD       QQENRKRHGS SRSVVDMDLD DTDDGDDNAP LFYQPGKRGF YTPRPGKNTE
HYD       QSSQPKK.. SPSVVVVDPV D....DDNEP LFYSPGKRGF YIPRQGFASF
C.ELEGANS QNAEQPRSDD EQWNVLVGTE SPLD.....Q LFEKAPYGEF YVPKIGNGAP

          2524                                                    2573
EDD       ARLNCFRNIG RILGLCLLQN ELCPITLNRH VIKVLLGRKV N:HDFAFFDP
HYD       ERINAFRNIG RLIGLCLLQN ELLPLFLQRH VLKYILRRKI K:HDLAFFDP
C.ELEGANS MRLAAYRTVG RIMGICLSQG DIFPLRFSRH IFSYILKLPI C:LDLGFYDP

          2574                                                    2623
EDD       VMYESLRQLI LASQSSDADA VFSAMDLAFA IDLCKEEGGG Q:ELIPNGVK
HYD       ALYESFRQII QNAQTKEGEE TINRMELCFV IDLMKEEGCG N:ELIPGG..
C.ELEGANS VLFNNLRALF .KAHPSELDT DFS......I VEQSADGKAV T:NLKPNGDE

          2624                                                    2671
EDD       RPVTPQNVYE YVRKDAEHRM LVV..AEQPL HAMRKGLLDV L:KNSLEDLT
HYD       RDVVSHRVIY SSTSDAIQNI DXIKSQEKAL EALKDGVFDV L:DNSMINLT
C.ELEGANS VLVNKDNVIE YVYKYAERVL ..VGKRIITF EAIREGILDV I:ESMLFGLT

          2672                                                    2721
EDD       AEDFRLLVNG CGEVNVQMLI SFTSFNDESG ENAEKLLQFK N:FWSIVEKM
HYD       AEDLRLLLNG VGDINVSTLI SYTTFNDESS EGPDKLXKFK K:FWSIVEKM
C.ELEGANS PEDLRLIICG IESVSISVLQ ENTGFLDESR ASQETLNRFK Q:FWQVIESF

          2722                                                    2771
EDD       SMTERQDLVY FWTSSPSLPA SEEGFQPMPS ITIRPPDDQH L:TANTCISR
HYD       NIMERQHLVY FWTGSPALPA SEEGFQPLPS VTIRPADDSH L:TANTCISR
C.ELEGANS TQQEKQELVF FWTGSPALPA AGK.WPTSAS VMLRPQEDVF L:TANTCISR

          2772
EDD       LYVPLYSSKQ ILKQKLLLAI KTKNFGFV
HYD       LYIPLYSSKS ILRSKMLMAI KSKNFGFV
C.ELEGANS LYVPVYSSKR VLRTKLLLAI KAKNFGFV
```

Figure 4 Sequence comparison between EDD and related proteins. (**a**) Similarity between central domains of EDD, HYD and a possible *C. elegans* homolog. Amino acid residues 1102 to 1493 of EDD are aligned with equivalent sequence from the HYD and *C. elegans* proteins.

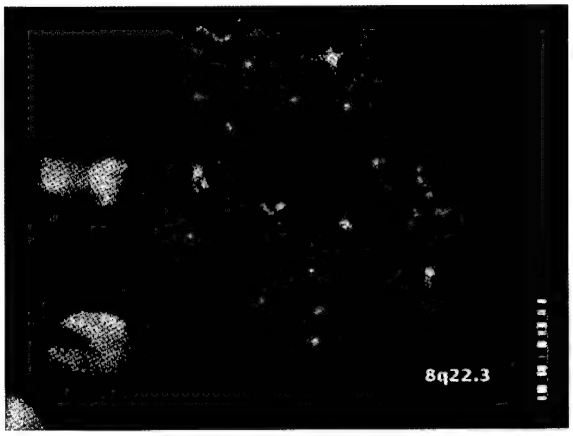


Figure 5 Chromosomal localization of the EDD gene. Metaphase showing FISH with the H1 probe. Normal male chromosomes were stained with DAPI. Hybridization sites on chromosome 8 are indicated by an arrow

We cloned a segment of rat cDNA containing the stretch of sequence upstream of the proposed initiation codon and found an additional base that, by changing the reading frame, removes the upstream stop codons (Figure 8). Correction of this apparent error results in a rat cDNA sequence that closely matches the human cDNA, in which a continuous open reading frame exists throughout the sequence. While the rat cDNA sequence corresponding to the amino terminal two-thirds of EDD has not been cloned, a number of mouse expressed sequences covering parts of this region are recorded in the GenBank database (Accession No. AA183561, AA177260, AA183970, AA231351, AA087561) and these show comparable levels of similarity with the EDD DNA sequence to that seen with the published rat sequence. Thus it appears that the true product of the rat gene is not a 100 kDa protein but may exist as a larger species. In rat lysates, however, AbPEP1 does not detect a protein having a molecular weight consistent with the human (EDD) and *Drosophila* (HYD) gene products.

Ubiquitin binding by EDD

A critical feature of the HECT family of E3 enzymes is their ability to reversibly form thioesters with ubiquitin at a conserved cysteine residue within the HECT domain. This property has been demonstrated for the HECT proteins human E6-AP, rat 100 kDa protein and yeast RSP5 where the thioester linkage remains intact in the absence of reducing agents but is broken in the presence of 100 mM DTT (Huibregtse *et al*., 1995). Substitution of the conserved cysteine residue prevents ubiquitin thioester bond formation. However, this property has not been shown for the HYD protein. To assess the potential of EDD to function as an E3 we tested whether EDD could form a reversible bond with ubiquitin via the conserved cysteine, C2768. ^{35}S-labeled *in vitro* translated truncated protein (~100 kDa of carboxyl terminus sequence) was incubated with purified GST-ubiquitin fusion protein in the presence or absence of DTT before SDS–PAGE (Figure 9a).

In the absence of DTT an additional higher molecular mass protein band was observed that corresponded to the expected size of an EDD-GST-ubiquitin conjugate (~130 kDa, upper arrow in Figure 9a). This species was abolished in the presence of 100 mM DTT suggesting involvement of a thioester bond in its formation. This was confirmed by experiments with an *in vitro* translated protein containing a C2768A mutation: binding of GST-ubiquitin was not seen under these conditions (Figure 9a). A species of slightly higher molecular mass than EDD was also observed (lower arrow in Figure 9a), consistent with the formation of ubiquitin-EDD conjugates, ubiquitin being present as a component of the rabbit reticulocyte lysate. Again this was not observed using the mutant protein or in the presence of 100 mM DTT. Similar results were achieved with full length EDD protein obtained (though at lower yield) by *in vitro* translation (Figure 9b).

Discussion

Application of the differential display PCR technique to a cultured human breast cancer cell model in which clearly defined proliferative responses to progestins are observed has led to the identification of a novel gene, EDD, a likely human homolog of the *Drosophila melanogaster* tumor suppressor gene *hyperplastic discs* (Mansfield *et al*., 1994). EDD is also highly homologous to the partial published sequence for the cDNA encoding the rat 100 kDa protein (Muller *et al*., 1992). All three genes produce large (approx 9.5 kb) mRNAs and the predicted entire EDD open reading frame of 2799 amino acids shares 40% identity with that of *Drosophila hyd* while the carboxy-terminal 889 amino acids of EDD share 96% identity with the rat protein. Western analysis showed that the EDD gene product is a protein of approximately 300 kDa. This protein is also immunoprecipitated by three different peptide-specific EDD antibodies and also corresponds to the size of the major *in vitro* translated gene product. The large discrepancy in the predicted size of the human and rat proteins was apparently resolved by re-examination of the rat cDNA sequence which disclosed an error in the published translation start site, pointing to the likelihood that a larger gene product exists.

At their carboxyl termini EDD, its rat homolog and HYD all contain a highly homologous HECT domain, indicating membership of a larger family of proteins which function as ubiquitin protein ligases (E3s). The

Amino acid residues conserved in at least two of the three proteins are indicated by bold type. Numbers refer to positions of amino acids in the EDD sequence. (b) Similarity between HECT domains of EDD, HYD and a possible *C. elegans* homolog. The carboxy-terminal 325 amino acids of EDD, including the HECT domain, are aligned with equivalent sequence from the HYD and C. elegans proteins. Amino acid residues conserved in at least two of the three proteins are indicated by bold type. Numbers refer to positions of amino acids in the EDD sequence

ubiquitination of target proteins occurs by the action of multiple interacting proteins: a ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzymes (E2) and ubiquitin-protein ligases (E3). Substrate specificity is largely determined by E3s, which bind and transfer ubiquitin to the target protein following interaction

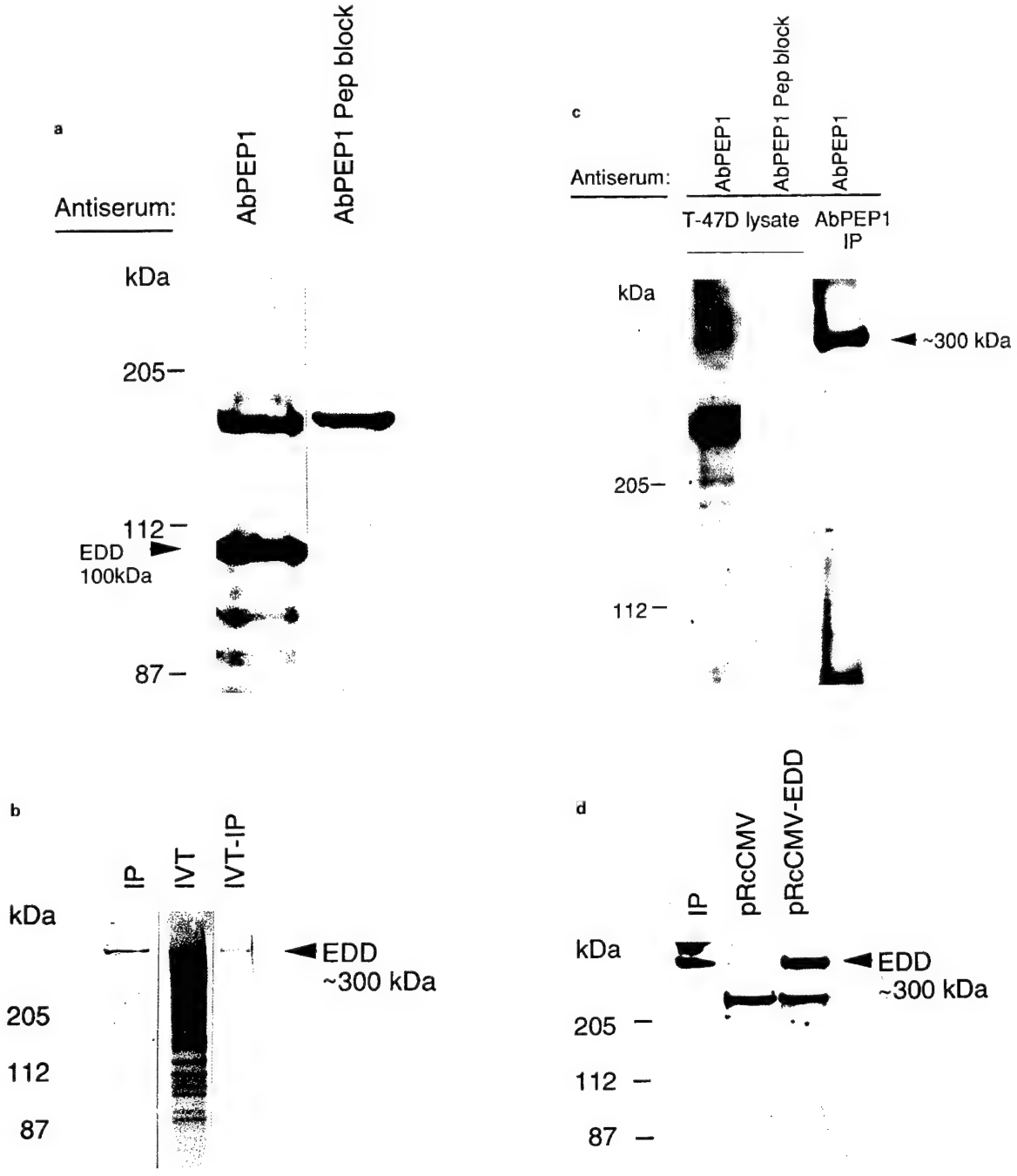


Figure 6 Characterization of EDD protein. (**a**) Detection of recombinant EDD protein with AbPEP1. *Sf9* cells infected with baculovirus containing a truncated EDD construct (EDD 100 kDa) were boiled in SDS-sample buffer prior to SDS–PAGE through a 6% gel, transferred to nitrocellulose and blotted with AbPEP1 or AbPEP1 peptide-blocked. (**b**) Determination of the size of the EDD protein. EDD was immunoprecipitated from T-47D lysate using AbPEP1. The immunoprecipitate (IP) was resolved by SDS–PAGE through a 6% gel alongside the products of *in vitro* translated full length EDD (IVT) and immunoprecipitated *in vitro* translated EDD (IVT-IP). The T-47D immunoprecipitate was transferred to nitrocellulose and blotted for EDD with AbPEP1 while the remainder of the gel was dried and autoradiographed. Molecular masses of marker proteins are indicated. (**c**) Detection of EDD protein in T-47D lysates. EDD immunoprecipitated from T-47D cell lysate was run alongside 40 μg lysate protein. Total proteins were blotted with either AbPEP1 or peptide-blocked AbPEP1 and the immunoprecipitate was blotted with AbPEP1. (**d**) Transient expression of recombinant EDD protein in HEK-293 cells. HEK-293 cells were transfected with either vector only (pRcCMV) or pRcCMV containing EDD cDNA (pRcCMV-EDD). Whole cell lysates (40 μg total protein) were run alongside EDD immunoprecipitated from T-47D whole cell lysate and blotted with AbPEP1

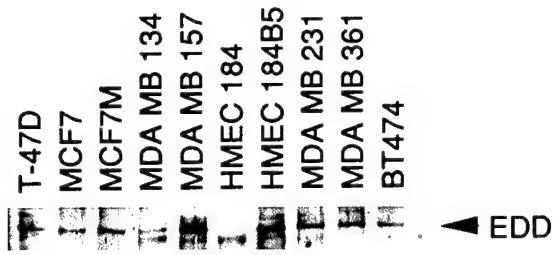


Figure 7 EDD protein expression in human tissues and cell lines. Expression of EDD in normal breast and breast cancer cell lines. Total cell lysates from a range of cell lines were separated by SDS–PAGE through a 6% gel, transferred to nitrocellulose and blotted with AbPEP1. 184 is a normal breast cell line, 184B5 an immortalized derivative, and the remainder are breast cancer cell lines, MCF-7M being a sub-line of MCF-7

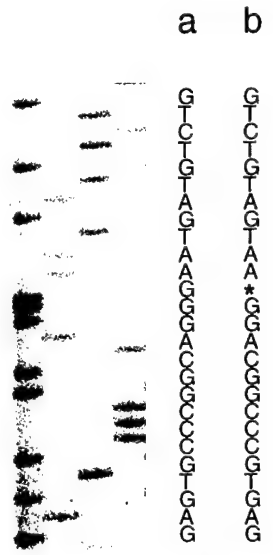


Figure 8 Sequence of the rat 100 kDa protein cDNA. Autoradiograph of the sequencing gel obtained when one clone was sequenced using the EDD-specific FC2 primer, with the sequence (a) listed alongside the autoradiograph. The published sequence (b (Muller *et al*., 1992)) is shown alongside and the missing base denoted by an asterisk

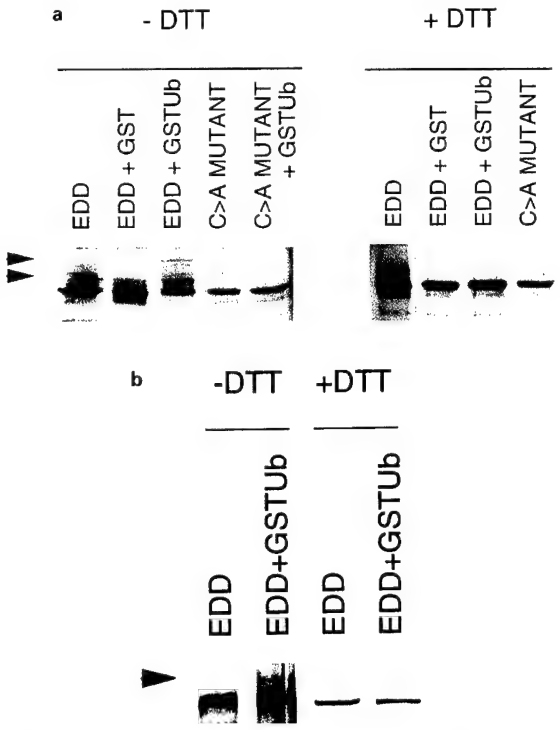


Figure 9 Ubiquitin thiol ester formation by EDD. *In vitro* translation of truncated (**a**) or full-length (**b**) EDD wild type or mutant (C2768A) protein in the presence of [^{35}S]methionine was followed by a 10 min incubation at 25°C either with or without purified GST-ubiquitin fusion protein (or GST in part a). Samples were resolved by SDS–PAGE (A, 7% gel; B, 6% gel) following either incubation at 25°C for 20 min in non-reducing sample buffer containing 4 M urea or boiling in sample buffer containing 100 mM DTT. Ubiquitin- and GST-ubiquitin-bound forms are marked with arrows

with specific E2s. The key feature of the HECT class of E3s is their ability to covalently bind ubiquitin through a conserved cysteine residue located in their HECT domains (Huibregtse *et al*., 1995). This property was demonstrated for EDD using *in vitro* translated protein that lost the ability to bind ubiquitin if the conserved cysteine (C2768) was substituted and we therefore conclude that EDD is an E3.

Few E3 genes have been cloned (only two from human) but others are likely to exist as ubiquitin-dependent proteolysis is involved in many cellular processes and targets many known proteins. Ubiquitin-mediated proteolysis is critical in the control of cell cycle progression, being responsible for the periodic destruction of key cell cycle regulators including cyclins (King *et al*., 1996; Won and Reed, 1996; Diehl *et al*., 1997) and cyclin-dependent kinase inhibitors (Pagano *et al*., 1995; Benito *et al*., 1998) and also targeting transcription factors (Palombella *et al*., 1994; Orian *et al*., 1995; Musti *et al*., 1996; 1997), the tumor suppressor protein p53 (Scheffner *et al*., 1993) and cell-cell signaling components such as β-catenin (Orford *et al*., 1997). Disruption of the murine Itch locus, which encodes an E3, caused hyperplasia in lymphoid and gastrointestinal epithelial tissues and an abnormal inflammatory response (Perry *et al*., 1998) while mutations in E6-AP in humans results in neurological abnormalities, indicating critical, and perhaps tissue specific, roles for E3 proteins (Kishino *et al*., 1997).

Although substrates for EDD and its rat and *Drosophila* homologs have yet to be defined, conservation between the central domain of EDD and that of HYD suggests that this region has an important function, perhaps in substrate recognition. For the yeast E3 Rsp5, substrate specificity is determined by the amino terminal domain and does not require the HECT domain (Huibregtse *et al*., 1997). Alternatively, this region could be involved in the binding of as yet unknown E2 proteins that interact specifically with EDD. The mouse E3 *Nedd4* has at least two distinct E2 binding domains, only one of which is within the HECT domain (Hatakeyama *et al*., 1997) while human E6-AP requires only the HECT domain for E2 recognition (Kumar *et al*., 1997b). As the protein produced from the truncated EDD construct still binds

ubiquitin reversibly, at least some E2 recognition function is present in this carboxyl domain. Other possible functions of the conserved central domain include cellular localization or translocation between cytoplasm and nucleus, cofactor association or phosphorylation.

Although ubiquitination is clearly involved in steroid-responsive processes such as regulation of cell cycle progression, specific regulation of ubiquitin pathways by steroid hormones has not previously been reported. The precise role of EDD in progestin action is unknown, particularly whether it participates in those key early events that occur in response to this hormone and which are ultimately responsible for its effects on cell proliferation and differentiation. Progestin regulation of EDD mRNA, which requires *de novo* protein synthesis, is transient with maximal levels three- to fourfold above control observed at 6 h. This increase in EDD expression levels therefore precedes the increase in the S phase fraction of T-47D cells following ORG 2058 treatment under the same conditions, which typically occurs at 12 to 14 h (Musgrove *et al*., 1991) and hence is consistent with a possible role in control of cell cycle progression. Similar levels of EDD induction were observed in antiestrogen-arrested MCF-7 breast cancer cells treated with 17β-estradiol (not shown), suggesting this may be a generalized response to mitogens.

However, given that EDD is also expressed in non progestin target tissues, a more widespread role than specifically mediating progestin effects is expected. Information on the biological role of HYD from mutagenesis studies in *Drosophila* (Mansfield *et al*., 1994) may ultimately give clues as to the function of EDD. The null *hyd* phenotype is lethal in the pupal or larval stages, as are severe mutations. Less severe mutations result in overgrowth (hyperplasia) of larval imaginal discs (the larval centers of cell proliferation that give rise to adult structures such as wings, legs and antennae), apparently caused by a failure to terminate cell proliferation when the discs reach their characteristic size, hence the definition of *hyd* as a tumor suppressor gene. Surviving adults are sterile due to germ cell defects, and interestingly, high expression levels of EDD and rat 100 kDa protein mRNA are seen in human and rat testes, suggesting a critical function in this organ.

Studies of a number of human homologs of *Drosophila* tumor suppressor genes strongly suggest that these genes have similar roles in both species in controlling cell proliferation, and that such genes can be important in human heritable and sporadic cancers, for example *patched* (Johnson *et al*., 1996), mutations of which are linked to basal cell carcinoma, and *discs large* (Kishino *et al*., 1997; Matsuura *et al*., 1997), a target of the APC gene which is mutated in sporadic colorectal tumors and familial adenomatous polyposis coli. The possible involvement of EDD in human tumorigenesis and tumor progression is therefore of particular interest. The EDD gene locus at chromosome 8q22 is often disrupted in a variety of cancers, being deleted in adenocarcinoma of the ovary and lung (Sato *et al*., 1994; Mitelman *et al*., 1997), hepatocellular carcinoma (Piao *et al*., 1998) and head and neck squamous cell carcinoma (Nawroz *et al*., 1994), amplified in many tumor types including gastrointestinal and primary breast cancers (Muleris *et al*., 1994; El-Rifai *et al*., 1996) and involved in translocations in acute myeloid leukemia (Erickson *et al*., 1992). Chromosome 8q22 is also a region affected in the human developmental disorder Klippel-Feil syndrome (Clarke *et al*., 1995).

In summary, EDD, a novel human gene, has been identified and appears to be a homolog of the *Drosophila* tumor suppressor gene *hyperplastic discs*. Its properties suggest the EDD protein is a member of the HECT domain family of ubiquitin-protein ligases. Major priorities for ongoing research are the identification of specific protein substrates for this enzyme and to determine whether the EDD gene has a tumor suppressor function that might be disrupted in human cancers.

Materials and methods

Reagents

Steroids and growth factors were obtained from the following sources: ORG 2058 (16α-ethyl-21-hydroxy-19-norpregn-4-en-3,20-dione), Amersham Australia Pty Ltd (Sydney, Australia); human transferrin (Sigma Chemical Co., St. Louis, MO, USA); and human insulin (Actrapid, CSL-Novo, North Rocks, Australia). Steroids were stored at $-20°C$ as 1000-fold-concentrated stock solutions in absolute ethanol. Cycloheximide (Calbiochem-Behring Corp., La Jolla, CA, USA) was dissolved at 20 mg/ml in water and filter sterilized. Actinomycin D (Cosmegen, Merck Sharp and Dohme Research Pharmaceuticals, Rahway, NJ, USA) was dissolved at 0.5 mg/ml in sterile water and used immediately. Tissue culture reagents were purchased from standard sources.

Cell culture

The sources and maintenance of the human breast cancer and normal cell lines used in this study were as described previously (Buckley *et al*., 1993; Hamilton *et al*., 1997), as were tissue culture experiments (Hamilton *et al*., 1997). Briefly, progestin (ORG 2058, 10 nM) and/or cycloheximide (20 μg/ml) or actinomycin D (5 μg/ml) was added to the medium and control flasks received the same volume of vehicle alone. To obtain RNA for differential display, cells were grown in insulin-supplemented serum-free medium and treated for 3 h with ORG 2058 or ethanol vehicle. Subsequent progestin stimulation experiments were carried out in medium containing 5% charcoal-stripped fetal calf serum without insulin.

RNA isolation and Northern analysis

Cells harvested from duplicate 150 cm^2 flasks were pooled, RNA extracted by a guanidinium-isothiocyanate-cesium chloride procedure and Northern analysis was performed as previously described with 20 μg of total RNA per lane (Alexander *et al*., 1989; Musgrove *et al*., 1991). The membranes were hybridized overnight (50°C) with probes labeled with [α-^{32}P]dCTP (Amersham Australia Pty Ltd) using a Prime-a-Gene labeling kit (Promega Corp., Sydney, Australia). The membranes were washed at a highest stringency of 0.2 × SSC (30 mM NaCl, 3 mM sodium citrate [pH 7.0])/1% sodium dodecyl sulfate at 65°C and exposed to Kodak X-OMAT or BIOMAX film at $-70°C$. Human multiple tissue Northern blots or RNA

Master blot (CLONTECH Laboratories Inc., Palo Alto, CA, USA) were hybridized under conditions recommended by the manufacturer. The mRNA abundance was quantitated by densitometric analysis of autoradiographs using Molecular Dynamics Densitometer and software (Molecular Dynamics, Sunnyvale, CA, USA). The accuracy of loading was estimated by re-hybridizing membranes with a [γ-^{32}P]ATP end-labeled oligonucleotide complementary to 18S rRNA (Chan *et al.*, 1984; Hall *et al.*, 1990).

Differential display

Differential display was carried out as described (Liang and Pardee, 1992) using a Heiroglyph mRNA Profile Kit No. 1 (Genomyx Corporation, Foster City, CA, USA) and recommended protocol. First strand cDNA synthesis was carried out in 96-well format 0.2 ml thin walled tubes. Typically 200 ng total RNA from T-47D cells treated with the synthetic progestin ORG 2058 for 3 h or from control T-47D cells was reverse transcribed with Expand Reverse Transcriptase enzyme (Boehringer Mannheim Pty Ltd, Castle Hill, Australia) following annealing with 4 pmol anchored primer (5′ACGACTCACTATAGGGCT$_{12}$AC). Subsequent PCR amplification was performed with one-tenth of the resultant cDNA in duplicate reactions containing [α-^{33}P]dATP with the anchored primer (0.2 μM), an arbitrary primer (5′ACAATTTCACACAGGAGCTAGCAGAC, 0.2 μM) and Expand Long Template Taq DNA Polymerase (Boehringer Mannheim). The PCR products were denatured and separated on a 4.5% denaturing polyacrylamide gel at 800 v for 16 h using the Genomyx Long Read Sequencing System reagents and apparatus. The gel was dried on the glass plate and exposed to X-ray film for 16–72 h. The DD-PCR product of interest was excised from the gel and amplified by PCR under the conditions recommended by the kit manufacturer using an M13 forward primer (5′AGCGGATAACAATTTCACACAGGA) and a T7 promoter primer (5′TAATACGACTCACTATAGGG). The reamplified PCR products were purified from 0.8% agarose gels using QIAEX reagents (Qiagen Pty Ltd, Clifton Hill, Australia).

Cloning and sequencing of cDNAs

Double stranded DNA templates were sequenced using the *fmol* DNA Cycle Sequencing System (Promega Corp.) with ^{33}P-labeled primers. The M13 primer was used for direct sequencing of DD-PCR products and the T7 and SP6 (5′GATTTAGGTGACACTATAG) promoter primers were used for sequencing PCR products cloned into the pGEM-T vector (Promega Corp.). Sequence database searches were performed at the NCBI using the Blast or Fasta network services. Peptide motif searches were carried out against the Prosite database.

Two primers (FC2: 5′GACGAAGGGCCCTGACTGCGCGAGAAGAAGC and R2: 5′AAAGAATTCTGTCATGGAGTCTGAACGTCG) that flank the region containing the reported rat 100 kDa start codon (Muller *et al.*, 1992) were used to amplify cDNA extracted from a rat hypothalamus library (CLONTECH). The resulting PCR product was cloned into pGEM-T (Promega Corp.) and four clones were sequenced.

Rapid amplification of cDNA 5′ ends (5′ RACE)

Additional sequence was obtained with the aid of a 5′RACE kit (Life Technologies Inc., Gaithersburg, MD, USA), following the manufacturer's instructions. Briefly, a gene specific primer (GSP1: 5′CACGCTCCAATGCAAGCTGG) was used to prime first strand cDNA synthesis. Following removal of the RNA strand, cDNA was 5′poly dC tailed and amplified by PCR. The target cDNA was amplified using an anchor primer (UAP: 5′GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG, where I represents deoxyinosine) in combination with a second gene specific primer (GSP2: 5′CGATCTTCCCTGATTCGAGGTGGC). Various gel-purified PCR products were further PCR amplified, primed by UAP and a third gene specific nested primer (GSP3: 5′CTGTATTGACAATGCTCCACC).

cDNA library screening

10^6 plaques from a human heart cDNA library in the Lambda ZAPII vector primed with both oligo (dT) and random primers (Stratagene, La Jolla, CA, USA) were transferred to nylon membranes (Hybond N, Amersham Australia Pty Ltd) and screened with both the original DD-PCR fragment and the RACE product as ^{32}P-labeled probes. This led to isolation of clone H1 (2.55 kb). This clone and the RACE product were used to screen 10^6 recombinants from a human placenta 5′STRETCH PLUS cDNA library in λgt10 primed with both oligo (dT) and random primers (CLONTECH Laboratories, Inc.). Sequencing of cDNA clones in either pBluescript or λgt10 was carried out as described above using vector- or gene-specific primers. Several rounds of isolation of positive clones and further screening of this library led to the isolation of the following overlapping clones covering the entire EDD open reading frame: P61 (1.95 kb), P43 (2.1 kb), P1 (1.5 kb), P19 (3 kb) and P47 (2.1 kb).

Fluorescence in situ *hybridization*

A probe corresponding to clone H1 was nick-translated with biotin-14-dATP and hybridized *in situ* at a final concentration of 20 ng/ml to metaphases from two normal males. The fluorescence *in situ* hybridization (FISH) method was modified from that previously described (Callen *et al.*, 1990) in that chromosomes were stained before analysis with both propidium iodide (as counterstain) and DAPI (for chromosome identification). Images of metaphase preparations were captured by a CCD camera using the CytoVision Ultra image collection and enhancement system (Applied Imaging Int Ltd). FISH signals and the DAPI banding pattern were merged for figure preparation.

Construction of recombinant cDNA clones for in vitro *translation and protein expression*

The full length EDD sequence was cloned by ligating three PCR products which spanned the open reading frame into pBluescript. The existing *Sal*I and *Eco*RI restriction sites used to ligate the fragments are indicated in Figure 3a. The carboxyl third of the cDNA was cloned into pBluescript such that a 890 amino acid truncated protein corresponding to the predicted rat 100 kDa protein (from aa 1910 to aa 2799) would be translated. An identical truncated cDNA fragment was cloned into the pFastBac1 expression vector (Life Technologies Inc.) for protein expression using the Bac-To-Bac baculovirus expression system in *Spodoptera frugiperda* (*Sf*9) cells and full length EDD cDNA was cloned into the pRcCMV expression vector (Invitrogen, Leek, The Netherlands) for transient transfection into HEK-293 cells. Mutagenesis of cysteine 2768 to alanine was performed for full length and truncated constructs in pBluescript using the Quick-Change site-directed mutagenesis kit (Stratagene). *In vitro* transcription and translation were performed using the TNT T7 Quick coupled rabbit reticulocyte lysate system (Promega Corp.) and [^{35}S]-methionine (1000 Ci/mmole, ICN Biomedicals Australasia Pty Ltd, Seven Hills, Australia).

SDS-polyacrylamide gel electrophoresis (PAGE) and immunoblotting

Cells growing in mid-log phase were lysed in 1% Triton X-100 buffer containing 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES; pH 7.5), 150 mM NaCl, 10% glycerol, 1.5 mM $MgCl_2$, 1 mM EGTA, 10 mM sodium pyrophosphate, 20 mM sodium fluoride, 1 mM dithiothreitol (DTT), 10 μg/ml each of aprotonin and leupeptin, 1 mM phenylmethylsulfonyl fluoride (PMSF) and 200 μM sodium orthovanadate. Lysates were cleared by centrifugation, quantitated according to a modified Bradford method (Bio-Rad Laboratories, Hercules, CA, USA) and typically 40 μg of total protein in SDS-sample buffer (50 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, and 0.2% bromophenol blue) containing 5% β-mercaptoethanol were resolved on 6% SDS-polyacrylamide gels. Following electrophoresis proteins were transferred to nitrocellulose (TransBlot, Bio-Rad Laboratories) and subjected to immunodetection. An EDD-specific peptide (SSEKVQQENRKRHGSS) was synthesized, coupled via glutaraldehyde to diptheria toxoid and used to generate a rabbit anti-EDD antibody (designated AbPEP1).

Immunoprecipitation

Cleared cell lysates (typically 1 mg total protein) or *in vitro* translation reactions were incubated with either control rabbit serum or AbPEP1 in the presence or absence of a tenfold excess of competing peptide for 1–2 h at 4°C. Following incubation with Protein A Sepharose 4B (Zymed, San Francisco, CA, USA), immunoprecipitates were washed three times in 1% Triton X100 lysis buffer described above, resolved by SDS–PAGE and either transferred to nitrocellulose and immunoblotted with AbPEP1 or where applicable dried onto Whatman 3 MM paper and subjected to autoradiography.

Ubiquitin-binding assay

^{35}S-labeled *in vitro* translated truncated (~100 kDa) or full length protein was tested for its ability to bind ubiquitin by incubating 5 μl translation reaction with or without 5 μg purified GST protein or GST-ubiquitin fusion protein for 10 min at 25°C (Scheffner *et al.*, 1995). Reactions were terminated by incubating the mixtures in either SDS-sample buffer containing 100 mM DTT at 95°C for 5 min or in SDS-sample buffer containing 4 M urea instead of DTT at 25°C for 20 min. Samples were resolved by SDS–PAGE through 6% or 7% gels followed by drying and autoradiography.

Acknowledgments

The authors thank Jenny Hamilton and Sharad Kumar for helpful discussions, Allan Shearn for the gift of the HYD antibody and Jon Huibregtse for the GST-ubiquitin plasmid. This work was supported by the National Health and Medical Research Council of Australia (NHMRC), the Kathleen Cuningham Foundation for Breast Cancer Research, the New South Wales State Cancer Council and the JH & JD Gunn Medical Research Foundation.

References

Alexander IE, Clarke CL, Shine J and Sutherland RL. (1989). *Mol. Endocrinol.*, **3,** 1377–1386.

Benito J, Martin-Castellanos C and Moreno S. (1998). *EMBO J.*, **17,** 482–497.

Buckley MF, Sweeney KJE, Hamilton JA, Sini RL, Manning DL, Nicholson RI, deFazio A, Watts CKW, Musgrove EA and Sutherland RL. (1993). *Oncogene,* **8,** 2127–2133.

Callen D, Baker E, Eyre HJ, Chernos JE, Bell JA and Sutherland GR. (1990). *Ann. Genet.*, **33,** 219–221.

Chan Y-L, Guttell R, Noller HF and Wool IG. (1984). *J. Biol. Chem.*, **259,** 224–230.

Clarke CL and Sutherland RL. (1990). *Endocr. Rev.*, **11,** 266–301.

Clarke RA, Singh S, McKenzie H, Kearsley JH and Yip MY. (1995). *Am. J. Hum. Genet.*, **57,** 1364–1370.

Dickson RB and Lippman ME. (1995). *Endocr. Rev.*, **16,** 559–589.

Diehl JA, Zindy F and Sherr CJ. (1997). *Genes. Dev.*, **11,** 957–972.

Dingwall C and Laskey RA. (1991). *Trends Biochem. Sci.*, **16,** 478–481.

El-Rifai W, Sarlomo-Rikala M, Miettinen M, Knuutila S and Andersson LC. (1996). *Cancer Res.*, **56,** 3230–3233.

Erickson P, Gao J, Chang KS, Look T, Whisenant E, Raimondi S, Lasher R, Trujillo J, Rowley J and Drabkin H. (1992). *Blood,* **80,** 1825–1831.

Groshong SD, Owen GI, Grimison B, Schauer IE, Todd MC, Langan TA, Sclafani RA, Lange CA and Horwitz KB. (1997). *Mol. Endocrinol.*, **11,** 1593–1607.

Gu J, Ren K, Dubner R and Iadarola MJ. (1994). *Mol. Brain Res.*, **24,** 77–88.

Hall RE, Lee CSL, Alexander IE, Shine J, Clarke CL and Sutherland RL. (1990). *Int. J. Cancer.*, **46,** 1081–1087.

Hamilton JA, Callaghan MJ, Sutherland RL and Watts CKW. (1997). *Mol. Endocrinol.*, **11,** 490–502.

Hatakeyama S, Jensen JP and Weissman AM. (1997). *J. Biol. Chem.*, **272,** 15085–15092.

Huibregtse JM, Scheffner M, Beaudenon S and Howley PM. (1995). *Proc. Natl. Acad. Sci. USA,* **92,** 5249.

Huibregtse JM, Scheffner M and Howley PM. (1993). *Mol. Cell. Biol.*, **13,** 775–784.

Huibregtse JM, Yang JC and Beaudenon SL. (1997). *Proc. Natl. Acad. Sci. USA,* **94,** 3656–3661.

Johnson RL, Rothman AL, Xie J, Goodrich LV, Bare JW, Bonifas JM, Quinn AG, Myers RM, Cox DR, Epstein E Jr and Scott MP. (1996). *Science,* **272,** 1668–1671.

King RW, Deshaies RJ, Peters JM and Kirschner MW. (1996). *Science,* **274,** 1652–1659.

Kishino T, Lalande M and Wagstaff J. (1997). *Nat. Genet.*, **15,** 70–73.

Kozak M. (1987). *Nucleic. Acids Res.*, **15,** 8125–8132.

Kumar S, Harvey KF, Kinoshita M, Copeland NG, Noda M and Jenkins NA. (1997a). *Genomics,* **40,** 435–443.

Kumar S, Kao WH and Howley PM. (1997b). *J. Biol. Chem.*, **272,** 13548–13554.

Laidlaw IJ, Clarke RB, Howell A, Owen AWMC, Potten CS and Anderson E. (1995). *Endocrinology,* **136,** 164–171.

Liang P and Pardee AB. (1992). *Science,* **257,** 967–971.

Lydon JP, DeMayo FJ, Funk CR, Mani SK, Hughes AR, Montgomery Jr CA, Shyamala G, Conneely OM and O'Malley BW. (1995). *Genes Dev.*, **9,** 2266–2278.

Mansfield E, Hersperger E, Biggs J and Shearn A. (1994). *Dev. Biol.*, **165,** 507–526.

Matsuura T, Sutcliffe JS, Fang P, Galjaard RJ, Jiang YH, Benton CS, Rommens JM and Beaudet AL. (1997). *Nat. Genet.*, **15,** 74–77.

Mitelman F, Mertens F and Johansson B. (1997). *Nat. Genet.*, **15,** 417–474.

Muleris M, Almeida A, Gerbault-Seureau M, Malfoy B and Dutrillaux B. (1994). *Genes Chromosomes Cancer,* **10,** 160–170.

Muller D, Rehbein M, Baumeister H and Richter D. (1992). *Nucleic. Acids Res.,* **20,** 1471–1475.

Musgrove EA, Hamilton JA, Lee CSL, Sweeney KJE, Watts CKW and Sutherland RL. (1993). *Mol. Cell. Biol.,* **13,** 3577–3587.

Musgrove EA, Lee CSL and Sutherland RL. (1991). *Mol. Cell. Biol.,* **11,** 5032–5043.

Musgrove EA, Swarbrick A, Lee CSL, Cornish AL and Sutherland RL. (1998). *Mol. Cell. Biol.,* **18,** 1812–1825.

Musti AM, Treier M and Bohmann D. (1997). *Science,* **275,** 400–402.

Musti AM, Treier M, Peverali FA and Bohmann D. (1996). *Biol. Chem.,* **377,** 619–624.

Nagase T, Miyajima N, Tanaka A, Sazuka T, Seki N, Sato S, Tabata S, Ishikawa K, Kawarabayasi Y, Kotani H and Nomura N. (1995). *DNA Res.,* **2,** 37–43.

Nawroz H, van der Riet P, Hruban RH, Koch W, Ruppert JM and Sidransky D. (1994). *Cancer Res.,* **54,** 11152–11155.

Nefsky B and Beach D. (1996). *EMBO J.,* **15,** 1301–1312.

Orford K, Crockett C, Jensen JP, Weissman AM and Byers SW. (1997). *J. Biol. Chem.,* **272,** 24735–24738.

Orian A, Whiteside S, Israel A, Stancovski I, Schwartz AL and Ciechanover A. (1995). *J. Biol. Chem.,* **270,** 21707–21714.

Pagano M, Tam SW, Theodoras AM, Beer-Romero P, Del Sal G, Chau V, Yew PR, Draetta GF and Rolfe M. (1995). *Science,* **269,** 682–685.

Palombella VJ, Rando OJ, Goldberg AL and Maniatis T. (1994). *Cell,* **78,** 773–785.

Perry WL, Hustad CM, Swing DA, O'Sullivan TN, Jenkins NA and Copeland NG. (1998). *Nat. Genet.,* **18,** 143–146.

Piao Z, Park C, Park J and Kim H. (1998). *Int. J. Cancer,* **75,** 29–33.

Prall OWJ, Sarcevic B, Musgrove EA, Watts CKW and Sutherland RL. (1997). *J. Biol. Chem.,* **272,** 10882–10894.

Sato S, Nakamura Y and Tsuchiya E. (1994). *Cancer Res.,* **54,** 5652–5655.

Scheffner M, Huibregtse JM, Vierstra RD and Howley PM. (1993). *Cell,* **75,** 495–505.

Scheffner M, Nuber U and Huibregtse JM. (1995). *Nature,* **373,** 81–83.

Staub O, Dho S, Henry P, Correa J, Ishikawa T, McGlade J and Rotin D. (1996). *EMBO J.,* **15,** 2371–2380.

Sutherland RL, Prall OWJ, Watts CKW and Musgrove EA. (1998). *J. Mamm. Gland Biol. Neoplasia,* **3,** 63–71.

van Gool AJ, Verhage R, Swagemakers SM, van de Putte P, Brouwer J, Troelstra C, Bootsma D and Hoeijmakers JH. (1994). *EMBO J.,* **13,** 5361–5369.

Won KA and Reed SI. (1996). *EMBO J.,* **15,** 4182–4193.